Part II REL Summary Notes

Tai Yingzhe, Tommy (ytt26)

Contents

1   Introduction

1.1   Units

Traditionally, you may be familiar with the widespread usage of the SI units. In most textbooks, the natural units are adopted, where we set the constants of nature as 1. This unifies physical dimensions and may highlight possible breakdowns of classical (Newtonian) theories.

Example 1.1 (Speed of Light). *From now on, $c = 1$. For speeds $v << c$, we recover Newtonian mechanics. But this becomes problematic at v close to 1, and we thus need Special Relativity.*

Example 1.2 (Gravitational Constant). *In addition, we further set $G = 1$. In that case, 1 s $= 3 \times 10^8$ m (from earlier) and 1 m $= 1.3466 \times 10^{27}$ kg. In this case, the Solar mass is roughly 1.47 km (This is the Schwarzschild radius of the Sun, as we will see later). For $\frac{M}{R} << \frac{c^2}{G} = 1$, Newtonian gravity is accurate. But for $\frac{M}{R} \approx 1$, we will require General Relativity.*

Example 1.3 (Planck's Constant). *While we do not need $\hbar$ in this course, we further demonstrate the idea of natural units by setting $\hbar = 1$. In that case, $\frac{\hbar}{mc} \approx R$ regime requires Quantum Mechanics, where $\frac{\hbar}{mc} = \frac{1}{m}$ is the Comptom wavelength.*

1.2   Newtonian Gravity

Proposition 1.1: For mutual collinear forces, the magnitude of the passive source is equal to that of the active source.

Proof. Follow from Newton's Third Law, together with the relevant force law. □

Proposition 1.2: Inertial mass is also the same as the active and passive mass.

Proof. Follows from Newton's Second Law. □

Remarks 1.1. *Proposition 0.1 is true for both masses and electric charges, but not true for Proposition 0.2 since there is no concept of inertial charge.*

Remarks 1.2. *The equality of gravitational mass (passive source which determines the gravitational force on the particle) and inertial mass has been experimentally verified to one part in 10^{13}. The many experiments that verified this, led to the equivalence principles.*

Proposition 1.3 (Weak Equivalence Principle (WEP)): Freely falling bodies with negligible gravitational self-interaction follow the same path if they have the same initial velocity and position, independent of composition.

Definition 1.1 (Local Inertial Frame): Local inertial frame is defined by a freely falling observer in the same way as an inertial frame is defined in Minkowski spacetime. Local means that the frame is much smaller than the length scale of gravitational field variations.

Einstein promoted the Weak Equivalence Principle.

Proposition 1.4 (Einstein Equivalence Principle): In a local inertial frame, the results of all non-gravitational experiments are indistinguishable from those of the same experiment performed in an inertial frame in Minkowski spacetime.

Proposition 1.5 (Strong Equivalence Principle (SEP)): In an arbitrary gravitational field, all the laws of physics (not just the dynamics of free-falling particles) in a free-falling, non-rotating laboratory occupying a sufficiently small region of spacetime look locally like special relativity (with no gravity).

Remarks 1.3. *In particular, the SEP implies that a constant gravitational field is unobservable – observations in a reference frame at rest in such a field would be indistinguishable from those in a uniformly-accelerating reference frame in the absence of gravity.*

Definition 1.2 (General Relativity): General relativity abandons the idea of gravity as a force defined on the fixed spacetime of special relativity, replacing it with a geometric theory in which the geometry of spacetime determines the trajectories of free-falling particles, the geometry itself being curved by the presence of matter.

Remarks 1.4.

1. *The SEP implies the WEP, but not the other way round.*

2. *In the SEP, we require the test body to be small, so that tidal effects are negligible.*

3. *The SEP is related to the equality of active and passive mass.*

4. *The SEP implies that Newton's gravitational constant G is the same everywhere in the universe, and suggests that gravity is entirely of geometrical nature. Otherwise, the gravitational binding energy of an extended object would depend on its position.*

5. *If the response of a body's motion to gravitational forces is independent of the properties of the body, it suggests that the gravitational force is not a feature of the body but exclusively of the spacetime in which it moves. To be more precise, gravity is a feature of the spacetime's geometry.*

1.3   Gravitational Waves

Definition 1.3 (Gravitational Waves)**:** Gravitational waves are wavelike disturbances in the geometry of spacetime, which can be detected by looking for their characteristic quadrupole distortion (i.e., a shortening in one direction and stretching in an orthogonal direction) of the two arms of a laser interferometer.

Gravitational waves propagate at the speed of light and are a natural prediction of general relativity; they do not arise in Newtonian gravity where the potential responds instantly to distant rearrangements of mass.

Example 1.4. *The first LIGO signal was generated by a truly extreme astrophysical source: two merging black holes each with a mass around 30 times that of the Sun at a distance from us of around 2 Gly.*

As the black holes orbited their common centre of mass, the system radiated gravitational waves causing the black holes to spiral inwards and increase their speed until they merged to form a single black hole.

Such sources probe the strong-field regime of general relativity during the merger phase and involve highly relativistic speeds. At its peak, the source was losing energy to gravitational waves at a rate of 3.6×10^{49} W, which is equivalent to 200 times the rest mass energy of the Sun per second!

2 Special Relativity Recap

We recap the concepts from IA Special Relativity.

2.1 Galilean Transformations

Definition 2.1 (Inertial frames)**:** Inertial frames are frames of references in which the frames themselves are not accelerating. Newton's Laws only hold in inertial frames.

Definition 2.2 (Galilean transformations)**:** Two inertial frames are related by

- Translations of space:
$$\mathbf{r}' = \mathbf{r} - \mathbf{r}_0$$

- Translations of time:
$$t' = t - t_0$$

- Rotation (and reflection):
$$\mathbf{r}' = R\mathbf{r}$$

with $R \in O(3)$.

They are simply symmetries of space itself.

Definition 2.3 (Galilean boost)**:** A Galilean boost is a change in frame of reference by

$$\mathbf{r}' = \mathbf{r} - \mathbf{v}t$$
$$t' = t$$

for a fixed, constant $\mathbf{v}$, i.e. uniform motion.

All these transformations together generate the Galilean group, which describes the symmetry of Newtonian equations of motion.

Law 2.1 (Galilean relativity)**:** The principle of relativity asserts that the laws of physics are the same in inertial frames.

Remarks 2.1. *The equations of Newtonian physics must have Galilean invariance. Since the laws of physics are the same regardless of your velocity, velocity must be a relative concept, and there is no such thing as an "absolute velocity" that all inertial frames agree on. However, all inertial frames must agree on whether you are accelerating or not (even though they need not agree on the direction of acceleration since you can rotate your frame). So acceleration is an absolute concept.*

Remarks 2.2 (Special Relativity)**.** *In special relativity, we abandon the notion of absolute time. It is replaced by a new postulate: The speed of light c is the same in all inertial frames.*

2.2 The Lorentz transformation

Proposition 2.1: Consider inertial frames S and S' whose origins coincide at $t = t' = 0$, and that S' moves with velocity v relative to S, then the Lorentz transformation must be of the form

$$x' = \gamma(x - vt)$$

for some factor γ.

Proof. For now, neglect the y and z directions, and consider the relationship between (x, t) and (x', t'). The general form is

$$x' = f(x, t), \quad t' = g(x, t)$$

for some functions f and g. In any inertial frame, a free particle moves with constant velocity. So straight lines in (x, t) must map into straight lines in (x', t'). Therefore the relationship must be linear. The line $x = vt$ must map into $x' = 0$, hence the result. We can use symmetry arguments to show that γ should take the same value for velocities v and $-v$). □

Remarks 2.3. *Now reverse the roles of the frames. From the perspective S', S moves with velocity $-v$. A similar argument leads to $x = \gamma(x' + vt')$ for the same γ, which is only dependent on $|v|$.*

Proposition 2.2: The factor γ is called the Lorentz factor and it is called

$$\gamma = \frac{1}{\sqrt{1 - v^2/c^2}}$$

Proof. Now consider a light ray (or photon) passing through the origin $x = x' = 0$ at $t = t' = 0$. Its trajectory in S is $x = c$. We want a γ such that the trajectory in S' is $x' = ct'$ as well, so that the speed of light is the same in each frame. Substitute these into the Lorentz transformation equations and multiply them together and finally divide by tt' to obtain

$$c^2 = \gamma^2(c^2 - v^2) \implies \gamma = \sqrt{\frac{c^2}{c^2 - v^2}} = \frac{1}{\sqrt{1 - (v/c)^2}}$$

$\square$

Proposition 2.3: The Lorentz transformation between times in two inertial frames is

$$t' = \gamma(t - vx/c^2)$$

Proof. Eliminate x between the previous two Lorentz transformation equations to obtain

$$x = \gamma(\gamma(x - vt) + vt') \implies t' = \gamma t - (1 - \gamma^{-2})\frac{\gamma x}{v} = \gamma\left(t - \frac{v}{c^2}x\right)$$

$\square$

Definition 2.4 (Standard Lorentz boosts)**:**

$$x' = \gamma(x - vt), \quad t' = \gamma(t - vx/c^2)$$

Remarks 2.4. *Directions perpendicular to the relative motion of the frames are unaffected.*

Remarks 2.5 (Generic Lorentz boosts)**.** *More generally, the relation between two Cartesian inertial frames S and S' can differ from that for the standard configuration since*

1. *the spacetime origins may not coincide, i.e. the event at $ct = x = y = z = 0$ may not be at $ct' = x' = y' = z' = 0$*

2. *the relative velocity of the two frames may be in an arbitrary direction in S, rather than along the x-axis*

3. *the spatial axes in S and S' may not be aligned*

The generic procedure is

1. *align the spacetime origins by appropriate temporal and spatial displacements*

2. *apply a purely spatial rotation in frame S to align the new x-axis with the relative velocity of the two frames*

3. *apply a standard Lorentz transform*

4. *apply a spatial rotation in the transformed coordinates to align the axes with those of S'.*

2.3 Spacetime diagrams

It is often helpful to plot out what is happening on a diagram. We plot them on a graph, where the position x is on the horizontal axis and the time ct is on the vertical axis. We use ct instead of t so that the dimensions make sense.

Definition 2.5 (Spacetime): The union of space and time in special relativity is called Minkowski spacetime. Each point P represents an event, labelled by coordinates (ct, x).

Definition 2.6 (World lines): A particle traces out a world line in spacetime, which is straight if the particle moves uniformly. Light rays moving in the x direction have world lines inclined at $45°$.

Remarks 2.6. *We can also draw the axes of S', moving in the x direction at velocity v relative to S. The ct' axis corresponds to $x' = 0$, i.e. $x = vt$. The x' axis corresponds to $t' = 0$, i.e. $t = vx/c^2$. Note that the x' and ct' axes are not orthogonal, but are symmetrical about the diagonal (dashed line). So they agree on where the world line of a light ray should lie on.*

Remarks 2.7. *For a massive particle passing through an event A, the particle's worldline must be inside the lightcone through A and each infinitesimal step must lie within the lightcone at each point. For a photon, the worldline will be tangent to the light cone.*

Definition 2.7 (Simultaneous events): We say two events P_1 and P_2 are simultaneous in the frame S if $t_1 = t_2$.

Remarks 2.8 (Causality). *Although different people may disagree on the temporal order of events, the consistent ordering of cause and effect can be ensured. Since things can only travel at at most the speed of light, P cannot affect R if R happens a millisecond after P but is at millions of galaxies away. We can draw a light cone that denotes the regions in which things can be influenced by P. These are the regions of space-time light (or any other particle) can possibly travel to. P can only influence events within its future light cone, and be influenced by events within its past light cone.*

Definition 2.8 (Invariant interval): The spacetime interval between P and Q is defined as

$$\Delta s^2 = c^2 \Delta t^2 - \Delta x^2$$

Note that this quantity Δs^2 can be both positive or negative.

Proposition 2.4: All inertial observers agree on the value of Δs^2.

Proof. Δx can be extended to $\mathbb{R}^3$:

$$
\begin{aligned}
c^2 \Delta t'^2 - \Delta x'^2 &= c^2 \gamma^2 \left(\Delta t - \frac{v}{c^2} \Delta x \right)^2 - \gamma^2 (\Delta x - v \Delta t)^2 \\
&= \gamma^2 \left(1 - \frac{v^2}{c^2} \right) (c^2 \Delta t - \Delta x^2) \\
&= c^2 \Delta t - \Delta x^2.
\end{aligned}
$$

$\square$

Remarks 2.9. *This allow us to calibrate the spacetime axes with respect to the invariant Δs^2.*

Definition 2.9 (Line element): The line element is

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2$$

Definition 2.10 (Timelike, spacelike and lightlike separation):

- Events with $\Delta s^2 > 0$ are timelike separated. It is possible to find inertial frames in which the two events occur in the same position, and are purely separated by time. Timelike-separated events lie within each other's light cones and can influence one another.

- Events with $\Delta s^2 < 0$ are spacelike separated. It is possible to find inertial frame in which the two events occur in the same time, and are purely separated by space. Spacelike-separated events lie out of each other's light cones and cannot influence one another.

- Events with $\Delta s^2 = 0$ are lightlike or null separated. In all inertial frames, the events lie on the boundary of each other's light cones. e.g. different points in the trajectory of a photon are lightlike separated, hence the name.

Definition 2.11 (4-vector)**:** A 4-vector is a four-component vector to describe an object's property in spacetime, in any inertial frame.

Definition 2.12 (4-Position)**:** The coordinates of an event P in frame S can be written as a 4-vector X. We write $X = (ct, \mathbf{x})^T$ where $\mathbf{x}$ is the 3-position, i.e. $\mathbf{x} \in \mathbb{R}^3$.

Proposition 2.5: The invariant interval between the origin and P can be written as an inner product. The inner product of any two 4-vectors on the Minkowski metric η is

$$X \cdot X = X^T \eta X = c^2 t^2 - x^2 - y^2 - z^2, \quad \eta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

4-vectors with $X \cdot X > 0$ are called timelike, and those $X \cdot X < 0$ are spacelike. If $X \cdot X = 0$, it is lightlike or null.

Definition 2.13 (Lorentz transformation)**:** A Lorentz transformation is a linear transformation of the coordinates from one frame S to another S', represented by a 4×4 tensor: $X' = \Lambda X$. Lorentz transformations can be defined as those that leave the inner product invariant, i.e. $\forall X$, $X' \cdot X' = X \cdot X$) which implies the matrix equation

$$\Lambda^T \eta \Lambda = \eta$$

These also preserve $X \cdot Y$ if X and Y are both 4-vectors.

Proposition 2.6: Two classes of solution to this equation are:

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & R & \\ 0 & & & \end{pmatrix}, \quad \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The former rotates space only and leaves time intact, where R is a 3×3 orthogonal matrix. For the latter, $\beta = \frac{v}{c}$, and $\gamma = 1/\sqrt{1 - \beta^2}$. Here we leave the y and z coordinates intact, and apply a Lorentz boost along the x direction.

Remarks 2.10. *The set of all matrices satisfying $\Lambda^T \eta \Lambda = \eta$ form the Lorentz group $O(1,3)$. It is generated by rotations and boosts, as defined above, which includes the absurd spatial reflections and time reversal. The subgroup with $\det \Lambda = +1$ is the proper Lorentz group $SO(1,3)$.*

The subgroup that preserves spatial orientation and the direction of time is the restricted Lorentz group $SO^+(1,3)$. Note that this is different from $SO(1,3)$, since if you do both spatial reflection and time reversal, the determinant of the matrix is still positive. We want to eliminate those as well!

Definition 2.14 (Rapidity)**:** The *rapidity* of a Lorentz boot is ϕ such that

$$\beta = \tanh \phi, \quad \gamma = \cosh \phi, \quad \gamma\beta = \sinh \phi.$$

Remarks 2.11. *Focus on the upper left 2×2 matrix of Lorentz boosts in the x direction. Write*

$$\Lambda[\beta] = \begin{pmatrix} \gamma & -\gamma\beta \\ -\gamma\beta & \gamma \end{pmatrix}, \quad \gamma = \frac{1}{\sqrt{1 - \beta^2}}$$

Combining two boosts in the x direction, we have

$$\Lambda[\beta_1]\Lambda[\beta_2] = \begin{pmatrix} \gamma_1 & -\gamma_1\beta_1 \\ -\gamma_1\beta_1 & \gamma_1 \end{pmatrix} \begin{pmatrix} \gamma_2 & -\gamma_2\beta_2 \\ -\gamma_2\beta_2 & \gamma_2 \end{pmatrix} = \Lambda \left[\frac{\beta_1 + \beta_2}{1 + \beta_1\beta_2} \right]$$

This is consistent with the idea of spatial rotations and thus add like rotation angles.

$$\Lambda[\beta] = \begin{pmatrix} \cosh \phi & -\sinh \phi \\ -\sinh \phi & \cosh \phi \end{pmatrix} = \Lambda(\phi), \quad \Lambda(\phi_1)\Lambda(\phi_2) = \Lambda(\phi_1 + \phi_2)$$

Lorentz boots are simply hyperbolic rotations in spacetime!

2.4 Relativistic kinematics

Definition 2.15 (Proper time)**:** The proper time τ is defined such that

$$\Delta\tau = \frac{\Delta s}{c}$$

τ is the time experienced by the particle, i.e. the time in the particles rest frame, the instantaneous rest frame (IRF) of the particle.

Remarks 2.12. *The world line of a massive particle can be parametrized using the proper time by $t(\tau)$ and $\mathbf{x}(\tau)$. Infinitesimal changes are related by*

$$d\tau = \frac{ds}{c} = \frac{1}{c}\sqrt{c^2\,dt^2 - |d\mathbf{x}|^2} = \sqrt{1 - \frac{|\mathbf{u}|^2}{c^2}}\,dt, \quad \frac{dt}{d\tau} = \gamma_u = \gamma_u = \frac{1}{\sqrt{1 - \frac{|\mathbf{u}|^2}{c^2}}}$$

The total time experienced by the particle along a segment of its world line is

$$T = \int d\tau = \int \frac{1}{\gamma_u}\,dt$$

Example 2.1 (Doppler Effect)**.** *Consider an observer $\mathcal{E}$ who moves at speed v along the x-axis of an inertial frame S in which an observer $\mathcal{O}$ is at rest at position x_0. Let successive wavecrests be emitted by $\mathcal{E}$ at events A and B, which are separated by proper time $\Delta\tau_{AB}$, i.e. this is the proper period of the source. The relation between $\Delta\tau_{AB}$ and the time between the emission events in S is*

$$\Delta\tau_{AB} = \sqrt{1 - \frac{v^2}{c^2}}\,\Delta t_e$$

The wavecrests are received by $\mathcal{O}$ at the events C and D, which are separated by time Δt_o in S; since $\mathcal{O}$ is at rest in S, the proper time between C and D is $\Delta\tau_{CD} = \Delta t_o$. In time Δt_e, the source $\mathcal{E}$ moves a distance $\Delta x_e = v\Delta t_e$ along the x-axis in S, and the second wavecrest has to travel Δx_e further than the first to be received by $\mathcal{O}$ at x_o. Then,

$$\Delta t_o = \left(1 + \frac{v}{c}\right)\Delta t_e$$

so that the ratio of proper times is

$$\frac{\Delta\tau_{AB}}{\Delta\tau_{CD}} = \frac{(1 - \beta^2)^{1/2}\Delta t_e}{(1 + \beta)\Delta t_e} = \sqrt{\frac{1 - \beta}{1 + \beta}}$$

This ratio is also the ratio of the received frequency, as measured by $\mathcal{O}$, to the proper frequency (i.e., the frequency in the rest-frame of the source $\mathcal{E}$).

Definition 2.16 (4-velocity)**:** Its 4-velocity is defined as

$$U = \frac{dX}{d\tau} = \begin{pmatrix} c\frac{dt}{d\tau} \\ \frac{d\mathbf{x}}{d\tau} \end{pmatrix} = \frac{dt}{d\tau}\begin{pmatrix} c \\ \mathbf{u} \end{pmatrix} = \gamma_u\begin{pmatrix} c \\ \mathbf{u} \end{pmatrix}, \quad \mathbf{u} = \frac{d\mathbf{x}}{dt}$$

Remarks 2.13. *U is a 4-vector because X is a 4-vector and τ is a Lorentz invariant. For any 4-vector U, the inner product $U \cdot U = U' \cdot U'$ is Lorentz invariant, i.e. the same in all inertial frames. In the rest frame of the particle, $U = (c, 0)$. So $U \cdot U = c^2$. In any other frame, $Y = \gamma_u(c, \mathbf{u})$. So*

$$Y \cdot Y = \gamma_u^2(c^2 - |\mathbf{u}|^2) = c^2$$

Proposition 2.7: A particle moves with constant velocity $u\hat{\mathbf{x}}$ in frame S. In a frame S' which moves with velocity $v\hat{\mathbf{x}}$ relative to S, the velocity is

$$u' = \frac{u - v}{1 - uv/c^2}$$

Proof. The world line of the particle in S' is $x' = u't'$. In S, using the inverse Lorentz transformation,

$$u = \frac{x}{t} = \frac{\gamma(x' + vt')}{\gamma(t' + (v/c^2)x')} = \frac{u't' + vt'}{t' + (v/c^2)u't'} = \frac{u' + v}{1 + u'v/c^2}$$

This is the formula for the relativistic composition of of velocities. The inverse transformation is found by swapping u and u', and swapping the sign of v, i.e.

$$u' = \frac{u - v}{1 - uv/c^2}$$

$\square$

Remarks 2.14. *If instead, the particle moves with constant velocity $u_x\hat{\mathbf{x}} + u_y\hat{\mathbf{y}}$ in frame S, then we have*

$$u'_y = \frac{u_y}{\gamma_v(1 - u_xv/c^2)}$$

Proposition 2.8: Consider a generic motion - not parallel to the Lorentz boost. In frame S, consider a particle moving with speed u at an angle θ to the x axis in the xy plane. In S', the composition of parallel velocities will give

$$u'\cos\theta' = \frac{u\cos\theta - v}{1 - \frac{uv}{c^2}\cos\theta}$$

while

$$\tan\theta' = \frac{u\sin\theta}{\gamma_v(u\cos\theta - v)},$$

which describes aberration, a change in the direction of motion of a particle due to the motion of the observer.

Proof. The 4-velocity is

$$U = \begin{pmatrix} \gamma_u c \\ \gamma_u u\cos\theta \\ \gamma_u u\sin\theta \\ 0 \end{pmatrix}, \quad \gamma_u = \frac{1}{\sqrt{1 - u^2/c^2}}$$

With frames S and S' in standard configuration (i.e. origin coincide at $t = 0$, S' moving in x direction with velocity v relative to S),

$$U' = \begin{pmatrix} \gamma_{u'}c \\ \gamma_{u'}u'\cos\theta' \\ \gamma_{u'}u'\sin\theta' \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma_v & -\gamma_v v/c & 0 & 0 \\ -\gamma_v v/c & \gamma_v & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma_u c \\ \gamma_u u\cos\theta \\ \gamma_u u\sin\theta \\ 0 \end{pmatrix}$$

Instead of evaluating the whole matrix, we can divide different rows to get the desired results. $\square$

Definition 2.17 (4-acceleration)**:** The 4-acceleration is defined as

$$A = \frac{dU}{d\tau}$$

where

$$U = \gamma_u \begin{pmatrix} c \\ \mathbf{u} \end{pmatrix} \implies A = \gamma_u\frac{dU}{dt} = \gamma_u \begin{pmatrix} \dot{\gamma}_u c \\ \gamma_u\mathbf{a} + \dot{\gamma}_u\mathbf{u}. \end{pmatrix}, \quad \mathbf{a} = \frac{d\mathbf{u}}{dt}, \ \dot{\gamma}_u = \gamma_u^3\frac{\mathbf{a}\cdot\mathbf{u}}{c^2}$$

Proposition 2.9: $U \cdot A = 0$ in all frames.

Proof. In the instantaneous rest frame of a particle, $\mathbf{u} = \mathbf{0}$ and $\gamma_u = 1$. So

$$U = \begin{pmatrix} c \\ \mathbf{0} \end{pmatrix}, \quad A = \begin{pmatrix} 0 \\ \mathbf{a} \end{pmatrix}$$

Then $U \cdot A = 0$. Since the inner product between any two 4-vector is a Lorentz invariant quantity, we have $U \cdot A = 0$ in all frames. $\square$

Remarks 2.15. *Acceleration is not invariant in special relativity, but is however, an absolute quantity in that all observers agree whether a particle is accelerating or not.*

Example 2.2 (Rectilinear acceleration)**.** *Consider a particle moving at a variable speed $u(\tau)$ along the x-axis in the inertial frame S, where τ is the particle's proper time. Let the particle carry an accelerometer that reads $f(\tau)$ – this is the proper acceleration, the acceleration in the instantaneous rest frame of the particle at τ.*

In the instantaneous rest frame at τ, $u'(\tau) = 0$ and $\frac{du'}{dt'} = f(\tau)$; transforming back to the frame S, we have

$$\frac{du}{dt} = \left(1 - \frac{u^2}{c^2}\right)^{3/2} f(\tau) = \left(1 - \frac{u^2}{c^2}\right) f(\tau)$$

In terms of rapidity $\psi(\tau)$, with $u(\tau) = c \tanh \psi(\tau)$, this is $cd\psi/d\tau = f(\tau)$, so by taking $u(\tau = 0) = 0$, we have

$$c\psi(\tau) = \int_0^\tau f(\tau')d\tau'$$

*To parameterise the worldline of the particle in S, we can use*

$$\frac{dt}{d\tau} = \gamma_u = \cosh \psi(\tau), \quad \frac{dx}{d\tau} = u\gamma_u = c \sinh \psi(\tau)$$

Integrating these equations gives the coordinates in S of the wordline, $t(\tau)$ and $x(\tau)$. Consider now the simple case of uniform or constant proper acceleration, i.e. $f = const$. The rapidity rises linearly with τ, i.e. $\psi(\tau) = f\tau/c$, and the worldline is

$$t = t_0 + \frac{c}{f} \sinh \frac{f\tau}{c}, \quad x = x_0 + \frac{c^2}{f}\left(\cosh \frac{f\tau}{c} - 1\right)$$

where t_0 and x_0 are integration constants. Setting $ct_0 = x_0 = 0$, we have a hyperbolic trajectory through the origin, as shown to the right, with an oblique asymptote in the future of $ct = c^2/f + x$. This means that there are regions of spacetime containing events that can never influence (i.e., communicate causally with) the accelerated particle (events to the left of the dotted line). The boundary of this region defines an event horizon of the accelerated observer.

As an example, light emitted from an object at rest at $x = 0$ in S will only reach the accelerated observer if it is emitted before $t = c/f$. Moreover, the accelerated observer sees the emitted light Doppler shifted to longer and longer wavelengths as the object approaches the event horizon and is observed as $\tau \to \infty$.
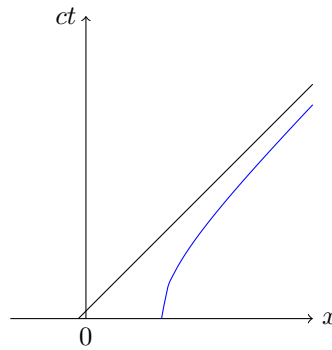
3   Differential Geometry

The spacetime of special relativity – Minkowski spacetime – is an example of what mathematicians call a manifold. In general relativity, spacetime is described by a more complicated manifold that responds dynamically to the distribution of mass and energy.

3.1   Manifolds and Coordinates

3.1.1   Manifolds and coordinate transformation

Definition 3.1 (Manifold)**:** An n-dimensional manifold $\mathcal{M}$ is a set of points with a bijective map to $\mathbb{R}^n$.

$$\phi : \ \mathcal{M} \to U \subset \mathbb{R}^n, \quad p \in \mathcal{M} \mapsto x^\alpha \in U \subset \mathbb{R}^n, \quad \alpha \in [0, n-1]$$

where U is an open subset of $\mathbb{R}^n$. x^α are the coordinates of point p on $\mathcal{M}$.

Remarks 3.1.

1. *Informally, an N-dimensional manifold is a set of objects that locally resembles N-dimensional Euclidean space $\mathbb{R}^N$.*

2. *It is not strictly required that we have one map ϕ that globally covers the entire manifold $\mathcal{M}$. It is, however, sufficient to partition $\mathcal{M}$ and map each part separately to $\mathbb{R}^n$. Everything we will develop also holds for such subdivisions of $\mathcal{M}$.*

3. *The manifold is differentiable if these subdivisions join up smoothly so that we can define scalar fields on the manifold that are differentiable everywhere.*

4. *Operations (taking derivatives) and objects (Curves, vectors etc). live on $\mathcal{M}$, not in the coordinate space U. But ϕ is one-to-one, so this distinction often blurred.*

5. *We can generally think of manifolds as surfaces embedded in some higher-dimensional Euclidean space, and we shall often do so, but it is important to appreciate that a given manifold exists independent of any embedding.*

6. *The coordinates are not unique: think of them as labels of points in the manifold that can change under a coordinate transformation (i.e., a change of map ϕ) while the point itself does not.*

Example 3.1.

1. *The Euclidean plane $\mathbb{R}^2$ is a 2D manifold that can be covered globally with the usual Cartesian coordinates. However, we could instead use plane-polar coordinates, (ρ, ϕ) with $0 \leq \rho < \infty$ and $0 \leq \phi < 2\pi$. Plane-polar coordinates are degenerate at $\rho = 0$ since ϕ is indeterminate there.*

2. *The 2-sphere $\mathcal{S}^2$ is the set of points in $\mathbb{R}^3$ with $x^2 + y^2 + z^2 = 1$. It is an example of a 2D manifold. The spherical polar coordinates (θ, ϕ), with $0 \leq \theta \leq \pi$ and $0 \leq \phi < 2\pi$, are degenerate at the poles $\theta = 0$ and $\theta = \pi$, where ϕ is indeterminate. For $\mathcal{S}^2$, there is no single coordinate system that covers the whole manifold without degeneracy: at least two coordinate patches are required.*

Definition 3.2 (Curve)**:** We define a curve to be a map $\lambda : I \subset \mathbb{R} \to \mathcal{M}$, where I is an open interval.

Remarks 3.2. *λ is smooth iff $\forall$ coordinate systems x^α, the map $x^\alpha \circ \lambda : I \to \mathbb{R}^n$ is smooth.*

Definition 3.3 (Submanifold)**:** A submanifold of $M < N$ dimensions is defined parametrically for some coordinate system

$$x^a = x^a(\{u^i\}), \quad a = 1, \ldots, N, \quad i = 1, \ldots M$$

Definition 3.4 (Hypersurface): A hypersurface is a submanifold with $M = N - 1$. Points in an M-dimensional submanifold can be specified by $N - M$ (independent) constraints.

$$f_i(\{x^j\}) = 0, \quad i = 1, \ldots, N - M, \quad j = 1, \ldots, M$$

Definition 3.5 (Coordinate transformations): We can assign new coordinates x'^a passively to a given point in terms of the original coordinates x^a. By chain rule,

$$dx'^a = \sum_{b=1}^{N} \frac{\partial x'^a}{\partial x^b} dx^b$$

This defines an $N \times N$ Jacobian matrix $J^\alpha{}_\beta = \frac{\partial x'^\alpha}{\partial x^\beta}$, where the index α labels the rows and the index β labels the columns.

Notation 3.1 (Einstein's summation convention). *Consider a sum $\mathbf{x} \cdot \mathbf{y} = \sum x_i y_i$. The summation convention says that we can drop the $\sum$ symbol and simply write $\mathbf{x} \cdot \mathbf{y} = x_i y_i$. If suffixes are repeated once, summation is understood. The rules of this convention are:*

1. *Suffix appears once in a term: free suffix*

2. *Suffix appears twice in a term: dummy suffix and is summed over*

3. *Suffixes cannot appear three times or more.*

Remarks 3.3. *Extend index notation:*

- *Distinguish upstairs and downstairs index (more later);*

- *Summation only over one up and one downstairs index, i.e. $v^j u_j := \sum_{j=1}^{3} v^j u_j$;*

- *Latin indices $i, j, \ldots = 1, 2, 3,$, Greek indices $\alpha, \beta, \ldots = 0 \ldots 3$*

Definition 3.6 (Riemannian manifold): In a Riemannian manifold $\mathcal{M}$, the local geometry near a point $p \in \mathcal{M}$ is specified by giving the invariant interval (independent of coordinate systems) between the points, which takes the form

$$ds^2 = g_{ab}(x) dx^a dx^b$$

where $g_{ab}(x)$ contain information about the local geometry but also depend on the particular coordinate system. g is called the metric (see later). Strictly, the geometry is Riemannian if $ds^2 > 0$ and pseudo-Riemannian otherwise (canbe positive, negative or zero).

Remarks 3.4.

1. *Without loss of generality, the metric $g_{ab}(x)$ is usually chosen to be symmetric. The antisymmetric contribution to ds^2 will vanish after relabelling the dummy indices $a \leftrightarrow b$.*

2. *It then follows that, in an N-dimensional Riemannian manifold, there are $0.5(N + 1)N$ independent metric functions at each point.*

Definition 3.7 (Intrinsic geometry): The interval ds^2 characterizes the local geometry (or curvature), which is an intrinsic property of the manifold independent of any possible embedding in some higher-dimensional space. Intrinsic properties are those that can be determined by an observer confined to the manifold.

Definition 3.8 (Global geometry or topology): Topology is an intrinsic, but non-local, property of a manifold.

Example 3.2.

1. *Consider the surface of a cylinder of radius a embedded in $\mathbb{R}^3$. In a cylindrical polar coordinate system (z, ϕ), the interval is $ds^2 = dz^2 + a^2 d\phi^2$. The intrinsic geometry is locally identical to the 2D Euclidean plane $\mathbb{R}^2$ by identifying $x \to a\phi$. However, the extrinsic geometry as seen within the embedding space $\mathbb{R}^3$ is clearly curved (non-Euclidean). However, the cylinder and the plane $\mathbb{R}^2$ have different topology.*

2. *The intrinsic geometry of a 2-sphere $\mathcal{S}^2$, however, is curved. It is not possible to find a coordinate transformation such that the sphere is transformed to Euclidean form over the entire surface.*

Example 3.3. *For a surface embedded in a higher-dimensional space, the induced line element in the surface is determined by the line element in the embedding space and the "shape" of the surface.*

1. *Consider the 2-sphere $\mathcal{S}^2$ embedded in $\mathbb{R}^3$; the embedding space has the Euclidean line element $ds^2 = dx^2 + dy^2 + dz^2$ in Cartesian coordinates. If the sphere has radius a, points on its surface satisfy $x^2 + y^2 + z^2 = a^2$, so that the constraint on dz gives*

$$0 = 2xdx + 2ydy = 2zdz \implies dz = -\frac{xdx + ydy}{\sqrt{a^2 - x^2 - y^2}} \implies ds^2 = dx^2 + dy^2 + \frac{(xdx + ydy)^2}{a^2 - x^2 - y^2}$$

Near the north or south poles, where $x^2 + y^2 << a^2$, the induced line element is approximately the Euclidean form, $ds^2 = dx^2 + dy^2$.

2. *Now consider the 3-sphere $\mathcal{S}^3$ $x^2 + y^2 + z^2 + w^2 = a^2$ embedded in $\mathbb{R}^4$. Similarly, the induced line element in spherical polar coordinates will be*

$$ds^2 = \frac{a^2}{a^2 - r^2}dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2$$

This describes the spatial part of a cosmological model with compact spatial sections, i.e. closed universe (see later). In the limit $a \to \infty$, we recover 3D Euclidean space in spherical polar coordinates. For $r << a$, we recover $\mathbb{R}^3$ locally.

3.1.2 Lengths and volumes

Definition 3.9 (Lengths along curves): Consider a curve $x^a(\lambda)$ between points A and B on some manifold, the invariant length along the curve is given by the invariant distance:

$$s = \int ds = \int_{\lambda_A}^{\lambda_B} \left| g_{ab} \frac{dx^a}{d\lambda} \frac{dx^b}{d\lambda} \right|^{1/2} d\lambda$$

where λ is a parameter along the curve.

Definition 3.10 (Volumes of regions): The invariant volume element of an N-dimensional manifold in an arbitrary coordinate system is

$$dV = \sqrt{|g|}dx^1 dx^2 \dots dx^N$$

Proposition 3.1: The volume element is indeed invariant.

Proof. Consider a coordinate transformation $x^a \to x'^a$, then

$$dx'^1 dx'^2 \dots dx'^N = J dx^1 dx^2 \dots dx^N$$

where $J = \det(\partial x'^a / \partial x^b)$ is the Jacobian. Since the metric transforms as

$$g'_{ab} = \frac{\partial x^c}{\partial x'^a} \frac{\partial x^d}{\partial x'^b} g_{cd}$$

then the determinant of the metric transforms as $g' = g/J^2$. It follows that

$$\sqrt{|g'|}dx'^1 dx'^2 \dots dx'^N = \frac{\sqrt{|g|}}{J} J dx^1 dx^2 \dots dx^N$$

and so $dV = \sqrt{|g|}dx^1 dx^2 \dots dx^N$ is indeed invariant. $\qquad\square$

Example 3.4. *Consider again $\mathcal{S}^2$ of radius a embedded in $\mathbb{R}^3$. We write the induced line element of the 2-sphere as $ds^2 = \frac{a^2 d\rho^2}{(a^2 - \rho^2)} + \rho^2 d\phi^2$. Consider the circle $\rho = R < a$. the distance form the centre O to the perimeter along the curve $\phi = const.$ is given by*

$$D = \int_0^R \frac{a}{\sqrt{a^2 - \rho^2}}d\rho = a \sin^{-1} \frac{R}{a}$$

The circumference of the circle is $2\pi R = 2\pi a \sin \frac{D}{a}$. The area enclosed is the 2-volume:

$$A = \int_0^{2\pi} \int_0^R \frac{a}{\sqrt{a^2 - \rho^2}} \rho d\rho d\phi = 2\pi a^2 \left[1 - \sqrt{1 - \frac{R^2}{a^2}} \right] = 2\pi a^2 \left[1 - \cos \frac{D}{a} \right]$$

For $D << a$, we recover the Eulicdean results $C = 2\pi D$ and $A = \pi D^2$. The coordinates (ρ, ϕ) are degenerate beyond the equator. However, if we switch to coordinates (D, ϕ), this system is well-defined beyond the equator and is only degenerate at $D = \pi a$ (the south pole).

3.1.3   Local coordinates and signature

Proposition 3.2: On a Riemannian manifold ($ds^2 > 0$) $\mathcal{M}$, it is not possible to choose coordinates such the line element takes the Euclidean form at every point. However, it is always possible to adopt coordinates such that in the neighbourhood of some point $p \in \mathcal{M}$, the line element takes the Euclidean form. More precisely, the local coordinates must satisfy

$$g_{ab}(p) = \delta_{ab}, \quad \left. \frac{\partial g_{ab}}{\partial x^c} \right|_p = 0$$

Proof. The first part follows since $g_{ab}(x)$ has $N(N+1)/2$ independent functions, but there are only N functions involved in coordinate transformations. Under the coordinate transformation, the metric and its derivatives transform as

$$g'_{ab} = \frac{\partial x^c}{\partial x'^a} \frac{\partial x^d}{\partial x'^b} g_{cd}$$

$$\frac{\partial g'_{ab}}{\partial x'^e} = \frac{\partial}{\partial x'^e} \left(\frac{\partial x^c}{\partial x'^a} \frac{\partial x^d}{\partial x'^b} \right) g_{cd} + \frac{\partial x^c}{\partial x'^a} \frac{\partial x^d}{\partial x'^b} \frac{\partial x^f}{\partial x'^e} \frac{\partial g_{cd}}{\partial x^f}$$

$\frac{\partial x^c}{\partial x'^a}|_p$ and $\frac{\partial^2 x^c}{\partial x'^e \partial x'^a}$ respectively have N^2 and $N^2(N+1)/2$ degrees of freedom. But, $g'_{ab}(p)$ and $\frac{\partial g'_{ab}}{\partial x'^c}|_p$ respectively have $N(N+1)/2$ and $N^2(N+1)/2$ degrees of freedom. The condition $g'_{ab}(p) = \delta_{ab}$ has $N(N+1)/2$ equations which have more equations than the N^2 degrees of freedom in $\frac{\partial x^c}{\partial x'^a}|_p$. On the other hand, the other condition $\frac{\partial g'_{ab}}{\partial x'^c}|_p = 0$ has $N^2(N+1)/2$ equations, just sufficient. □

Remarks 3.5.

1. *The remaining $-N(N+1)/2 + N^2 = (-N + N^2)/2$ equations for $N = 4$ correspond to 3 Lorentz boost and 3 rotations (total 6).*

2. *We cannot set the second derivatives of the metric to be zero. The condition $\frac{\partial^2 g'_{ab}}{\partial x'^c \partial x'^d} = 0$ has $N^2(N+1)^2/4$ equations, but $\frac{\partial^3 x^a}{\partial x'^b \partial x'^c \partial x'^d}$ has $N^2(N+1)(N+2)/6$ degrees of freedom. There are thus $0.25N^2(N+1)^2 - N^2(N+1)(N+2)(1/6) = N^2(N^2-1)/12$ independent degree of freedom that cannot be eliminated, and corresponds to the curvature of the manifold. For $N = 4$, there are 20 of them, associated with gravity.*

Proposition 3.3: In a pseudo-Riemannian manifold $\mathcal{M}$, one can always find coordinates such that at a point $p \in \mathcal{M}$,

$$g_{ab}(p) = \text{diag}(\pm 1, \pm 1, \dots, \pm 1), \quad \left. \frac{\partial g_{ab}}{\partial x^c} \right|_p = 0$$

Theorem 3.1 (Sylvester's Law): The number of $+1$ and -1 entries in the metric is independent of the basis.

Definition 3.11 (Signature): The signature σ of a metric $g_{\alpha\beta}$ on an n-dimensional manifold $\mathcal{M}$, is the sum of $+1$ and -1 over all diagonal elements. The signature is usually convention dependent.

3.2 Vector and Tensor Algebra

3.2.1 Vectors and tensors

Definition 3.12 (Functions on Manifold)**:** A function on manifold is a mapping $f : \mathcal{M} \to \mathbb{R}$. If f is invariant under a change of coordinates, it is a scalar.

Remarks 3.6. *f is smooth iff $\forall$ coordinate systems x^α, $f(x^\alpha)$ is a smooth function $f : \mathbb{R}^n \to \mathbb{R}$.*

Definition 3.13 (Vectors)**:** Let C^∞ be the space of all smooth functions $f : \mathcal{M} \to \mathbb{R}$, λ be a smooth curve and $p = \lambda(0) \in \mathcal{M}$. The tangent vector to the curve λ at $p \in \mathcal{M}$ is the map

$$V : \ C^\infty \to \mathbb{R}, \quad f \mapsto V(f) = \frac{d}{dt} f(\lambda(t)) \Big|_{t=0}$$

Remarks 3.7. *In Euclidean space, displacement vectors connect two points in the space while local vectors are measured at a given observation point and refer solely to that point. Displacement vectors between infinitesimally separated points are really local vectors, as are derivatives of displacement vectors. On a general manifold, we can only define local vectors. Displacement vectors do make sense if we specify an embedding of $\mathcal{M}$ in some higher-dimensional Euclidean space, but we are interested only in intrinsic geometry here.*

Definition 3.14 (Tangent Space)**:** $\mathcal{T}_p(\mathcal{M})$ is the tangent space of all vectors at $p \in \mathcal{M}$.
Essentially, a vector is a derivative operator. It obeys

- linearity:
$$V(\alpha f + \beta g) = \alpha V(f) + \beta V(g), \quad \alpha, \beta \in \mathbb{R}, \quad f, g \in C^\infty$$

- Leibniz rule:
$$V(fg) = V(f)g + V(g)f, \quad f, g \in C^\infty$$

Consider coordinate system x^α on the manifold $\mathcal{M}$,

$$V(f) = \frac{d}{dt} f(x^\mu(\lambda(t))) = \frac{dx^\mu}{dt} \Big|_\lambda \frac{\partial}{\partial x^\mu} f(x^\alpha)$$

where the components are $V^\mu := \frac{dx^\mu}{dt}|_\lambda$ and bases are $e_\mu := \partial_\mu := \frac{\partial}{\partial x^\mu}$, and are called coordinate basis.

Proposition 3.4: $\mathcal{T}_p(\mathcal{M})$ has dimension n and $\frac{\partial}{\partial x^\mu}$ form a basis of $\mathcal{T}_p(\mathcal{M})$.

Proposition 3.5: Under a coordinate change $x^\mu \to \tilde{x}^\alpha$, the components and bases of a vector transform like

$$e_\mu \mapsto \tilde{e}_\alpha = \frac{\partial}{\partial \tilde{x}^\alpha} = \frac{\partial x^\mu}{\partial \tilde{x}^\alpha} e_\mu$$

$$V^\mu \mapsto \tilde{V}^\alpha = \frac{\partial \tilde{x}^\alpha}{\partial x^\mu} V^\mu$$

so $V = V^\mu e_\mu = \tilde{V}^\alpha \tilde{e}_\alpha$ is invariant under this coordinate change.

Proof. Chain Rule. $\qquad\square$

Remarks 3.8. *The components and the basis depend on the coordinates, but the vector is a geometric object, invariant under coordinate transformation.*

Definition 3.15 (Covector)**:** Covectors, also known as one-forms or dual vectors, are linear map

$$\eta : \ \mathcal{T}_p(\mathcal{M}) \to \mathbb{R}, \quad V \mapsto \eta(V)$$

Definition 3.16 (Co-tangent Space)**:** $\mathcal{T}_p^*(\mathcal{M})$ is the co-tangent space of all co-vectors at $p \in \mathcal{M}$. $\mathcal{T}_p^*(\mathcal{M})$ is also an n-dimensional vector space. Let e_μ be a basis of $\mathcal{T}_p(\mathcal{M})$. The components of a covector η are $\eta_\mu := \eta(e_\mu)$. They satisfy

- Linearity:
$$\eta(\alpha V + \beta W) = \alpha \eta(V) + \beta \eta(W), \quad \alpha, \beta \in \mathbb{R}, \quad V, W \in \mathcal{T}_p(\mathcal{M})$$

- Components:

$$\eta(V) = \eta(V^\mu e_\mu) = V^\mu \eta_\mu$$

Again, require co-vectors to be invariant under coordinate transformations, since they are scalars.

Proposition 3.6: Under a coordinate change $x^\mu \to \tilde{x}^\alpha$, the components of a co-vector transform like

$$\tilde{\eta}_\beta = \frac{\partial x^\mu}{\partial \tilde{x}^\beta} \eta_\mu$$

Proof. Since one-forms are scalars, result follow from chain rule. $\qquad \square$

Definition 3.17 (Gradient)**:** Gradient $\mathbf{d}f$ of a smooth function f.

$$\mathbf{d}f : \ \mathcal{T}_p(\mathcal{M}) \to \mathbb{R}, \quad V = \frac{d}{dt} \mapsto \frac{df}{dt} = V(f), \quad V \in \mathcal{T}_p(\mathcal{M})$$

Notation 3.2. *We bold d for notational purposes, to differentiate df and $\mathbf{d}f$.*

Proposition 3.7: The gradient form a basis for co-vectors, and are dual basis of the vector basis ∂_μ.

Proof. Let $f = x^\alpha$ with α fixed, then the gradient is

$$\mathbf{d}x^\alpha(e_\beta) = \mathbf{d}x^\alpha \frac{\partial}{\partial x^\beta} = \delta^\alpha{}_\beta$$

then,

$$\eta_\alpha \mathbf{d}x^\alpha(V^\beta \partial_\beta) = \eta_\alpha V^\beta \mathbf{d}x^\alpha(\partial_\beta) = \eta_\alpha V^\beta \delta^\alpha{}_\beta = \eta_\alpha V^\alpha = \eta(V)$$

$$\square$$

Definition 3.18 (Tensors)**:** A tensor T at $p \in \mathcal{M}$ of rank (r, s) with $r, s \in \mathbb{N} \cup \{0\}$ is a multilinear map.

$$T : (\mathcal{T}_p^*(\mathcal{M}))^r \times (\mathcal{T}_p(\mathcal{M}))^s \to \mathbb{R}$$

where we take Cartesian product of r number of co-tangent spaces and s number of tangent spaces.

Example 3.5. *We first consider simple examples of tensors:*

- *Covector η is a tensor of rank (0,1).*

- *Vector is a tensor of rank (1,0) and can be viewed as $V : \mathcal{T}_p^*(\mathcal{M}) \to \mathbb{R}$, $\eta \to \eta(V)$. The components of V are*

$$\eta(V) = \eta_\alpha \mathbf{d}x^\alpha(V) = n_\alpha V^\alpha \implies V^\alpha = \mathbf{d}x^\alpha(V) = V(\mathbf{d}x^\alpha)$$

- *Kronecker-Delta tensor*

$$\delta : \ \mathcal{T}_p^*(\mathcal{M}) \times \mathcal{T}_p(\mathcal{M}) \to \mathbb{R}, \quad (\eta, V) \mapsto \eta(V)$$

$\forall \eta \in T_p^*(\mathcal{M})$, $V \in T_p(\mathcal{M})$ and is a tensor of rank (1,1) with components

$$\delta(\mathbf{d}x^\alpha, \partial_\beta) = \mathbf{d}x^\alpha(\partial_\beta) = \frac{\partial x^\alpha}{\partial x^\beta} = \delta^\alpha{}_\beta$$

In general, for all tensors:

$$T^{\alpha_1 \dots \alpha_r}{}_{\beta_1 \dots \beta_s} = T(\mathbf{d}x^{\alpha_1}, \dots, \mathbf{d}x^{\alpha_r}, e_{\beta_1}, \dots, e_{\beta_s})$$

Proposition 3.8: Tensors transform like

$$\tilde{T}^{\alpha_1 \dots \alpha_r}{}_{\beta_1 \dots \beta_s} = \frac{\partial \tilde{x}^{\alpha_1}}{\partial x^{\mu_1}} \dots \frac{\partial \tilde{x}^{\alpha_r}}{\partial x^{\mu_r}} \frac{\partial x^{\nu_1}}{\partial \tilde{x}^{\beta_1}} \dots \frac{\partial x^{\nu_s}}{\partial \tilde{x}^{\beta_s}} T^{\mu_1 \dots \mu_r}{}_{\nu_1 \dots \nu_s}$$

Proof. Chain Rule for $r + s$ number of times. $\qquad\square$

Proposition 3.9: The proper distance in this metric is $ds^2 = g_{\alpha\beta}dx^\alpha dx^\beta$ and is invariant under a change of coordinates.

Proof.

$$d\tilde{s}^2 = \tilde{g}_{\mu\nu}d\tilde{x}^\mu d\tilde{x}^\nu = \frac{\partial x^\alpha}{\partial \tilde{x}^\mu}\frac{\partial x^\beta}{\partial \tilde{x}^\nu}g_{\alpha\beta}\frac{\partial \tilde{x}^\mu}{\partial x^\rho}\frac{\partial \tilde{x}^\nu}{\partial x^\sigma}dx^\rho dx^\sigma = ds^2$$

$\qquad\square$

Tensor operations:

1. Addition and scalar multiplication: For instance,

$$c_1 S + c_2 T : \mathcal{T}_p^*(\mathcal{M}) \times \mathcal{T}_p(\mathcal{M}) \to \mathbb{R}, \quad \eta, V \mapsto c_1 S(\eta, V) + c_2 T(\eta, V)$$

   where $c_1, c_2 \in \mathbb{R}$ and S, T are (1,1) tensors.

2. Symmetrization/anti-symmetrization:

   - Consider a (0,2) tensor T, the symmetric part and anti-symmetric part are respectively

$$S_{\alpha\beta} := \frac{1}{2}(T_{\alpha\beta} + T_{\beta\alpha}), \quad A_{\alpha\beta} := \frac{1}{2}(T_{\alpha\beta} - T_{\beta\alpha})$$

   They can also be denoted as $T_{(\alpha\beta)}$ and $T_{[\alpha\beta]}$ respectively.
   - Index subsetting:
$$T^{(\alpha\beta)\gamma}{}_\delta := \frac{1}{2}(T_\delta^{\alpha\beta\gamma} + T_\delta^{\beta\alpha\gamma})$$

   - Non-adjacent indices:
$$T_{(\alpha|\beta\gamma|\delta)} := \frac{1}{2}(T_{\alpha\beta\gamma\delta} - T_{\delta\beta\gamma\alpha})$$

   - Anti-symmetrize two indices. For $n > 2$ indices, we sum over all permutations, apply the sign of permutation for this anti-symmetrization operation and finally divided by $n!$. For instance,

$$T^\alpha{}_{[\beta\gamma\delta]} = \frac{1}{3!}(T^\alpha{}_{\beta\gamma\delta} + T^\alpha{}_{\delta\beta\gamma} + T^\alpha{}_{\gamma\delta\beta} - T^\alpha{}_{\delta\gamma\beta} - T^\alpha{}_{\beta\delta\gamma})$$

3. Contraction of (r, s) tensor: summation over 1 upper and 1 lower index to give $(r-1, s-1)$ tensor. For instance, let T be a (3,2) tensor $T^{\alpha\beta\gamma}{}_{\alpha\delta}$ and we contract it to yield a (2,1) tensor $S^{\beta\gamma}{}_\delta$ $S(\omega, \eta, V) := T(\mathbf{d}x^\mu, \omega, \eta, \partial_\mu, V)$. This is basis independent:

$$T(\mathbf{d}\tilde{x}^\mu, \omega, \eta\frac{\partial}{\partial \tilde{x}^\mu}, V) = \frac{\partial \tilde{x}^\mu}{\partial x^\alpha}\frac{\partial x^\beta}{\partial \tilde{x}^\mu}T(\mathbf{d}x^\alpha, \omega, \eta, \partial_\beta, V) = \delta^\beta{}_\alpha T(\mathbf{d}x^\alpha, \omega, \eta, \partial_\beta, V) = T(\mathbf{d}x^\alpha, \omega, \eta, \partial_\alpha, V)$$

4. Outer product: Let S be a (p, q) tensor and T be a (r, s) tensor, then $S \otimes T$ is a $(p+r, q+s)$ tensor. We have

$$(S \otimes T)(\omega_1, ..., \omega_p, \eta_1, ..., \eta_r, X_1, .., X_q, Y_1, ..., Y_s) := S(\omega_1, ..., \omega_p, X_1, ..., X_q)T(\eta_1, ..., \eta_r, Y_1, ..., Y_s)$$

   In another words,

$$(S \otimes T)^{\alpha_1...\alpha_p\beta_1...\beta_r}{}_{\mu_1...\mu_q\nu_1...\nu_s} = S^{\alpha_1...\alpha_p}{}_{\mu_1...\mu_q}T^{\beta_1...\beta_r}{}_{\mu_1...\mu_s}$$

   For instance, in a coordinate basis, a (2,1) tensor can be written as

$$T = T^{\mu\nu}{}_\rho e_\mu \otimes e_\nu \otimes \mathbf{d}x^\rho$$

Remarks 3.9 (Quotient theorem). *If a set of quantities when contracted with an arbitrary tensor produces another tensor, the original set of quantities form the components of a tensor.*

Definition 3.19 (Tensor Fields)**:** A tensor field of rank (r,s) is a collection of (r,s) tensors at each point $p \in \mathcal{M}$. Like a map $p \mapsto T_p$ of rank (r,s).

Remarks 3.10. *The tensor field is smooth iff its component functions in a coordinate basis are smooth functions.*

Sometimes we write X_p being a vector and X being a field.

Example 3.6. *A vector field $X : \mathcal{M} \to \mathcal{T}_p(\mathcal{M}), p \mapsto X_p$. A function f $X(f) : \mathcal{M} \to \mathbb{R}$, $p \mapsto X_p(f)$.*

Henceforth we assume all tensor fields to be smooth.

3.2.2 Metric Tensor

Now, let's properly discuss the metric tensor:

Definition 3.20 (Metric)**:** A metric at $p \in \mathcal{M}$ is a (0,2) tensor that is

- symmetric: $g(V,W) = g(W,V)$ $\forall V,W \in \mathcal{T}_p(\mathcal{M})$. Equivalently, $g_{\alpha\beta} = g_{\beta\alpha}$.

- non-degenerate: $g(V,W) = 0$ $\forall W \in \mathcal{T}_p(\mathcal{M})$. Equivalently, $V = 0$.

The components of a metric are $g = g_{\alpha\beta} \mathbf{d}x^\alpha \otimes \mathbf{d}x^\beta \implies g_{\mu\nu} = g(\partial_\mu, \partial_\nu)$.

A metric maps vectors to one-forms. $g : V \mapsto \underline{V} := g(V,.)$ where $\underline{V}$ is a map

$$\underline{V} : \mathcal{T}_p(\mathcal{M}) \to \mathbb{R}, \quad W \mapsto \underline{V}(W) := g(V,W) = \underline{V}_\mu W^\mu = g_{\mu\nu} V^\mu W^\nu$$

The components are thus $V_\mu := \underline{V}_\mu = g_{\mu\nu} V^\nu$.

Theorem 3.2: If g is non-degenerate, then g is invertible.

Definition 3.21 (Inverse Metric)**:** We define the inverse metric to be g^{-1} such that it is a symmetric (2,0) tensor, i.e. $(g^{-1})^{\alpha\beta} g_{\beta\gamma} = \delta^\alpha{}_\gamma$. Now, g^{-1} maps one-forms to vectors, i.e. $g^{-1} : \underline{V} \mapsto V$.

Remarks 3.11. *From now on, we will drop the exponent -1 when we write the components of the inverse metric and merely distinguish it from the metric by the position of the indices.*

Example 3.7. *The line element on the unit sphere, $x^2 + y^2 + z^2 = 1$ in $\mathbb{R}^3$ is $ds^2 = d\theta^2 + \sin^2\theta d\phi^2$. The metric is*

$$g_{\alpha\beta} = \begin{bmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{bmatrix}, \quad g^{\alpha\beta} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\sin^2\theta} \end{bmatrix}$$

Theorem 3.3: $\exists$ a natural isomorphism between vectors and one-forms, i.e. $g^{-1}(g(V,.),.) = V$ and $g(g^{-1}(\eta,.),.) = \eta$.

Proposition 3.10: The following statements on the metric tensor are true:

- $\exists$ a basis where $g_{\mu\nu}$ is diagonal;

- $\exists$ a non-unique orthonormal basis for g.

Proof. By construction, g symmetric and non-degenerate. It being symmetric $\implies$ components of g at $p \in \mathcal{M}$ are a symemtric basis, hence there exists a basis where $g_{\mu\nu}$ is diagonal. It being non-degenerate $\implies$ all eigenvalues $\neq 0$ $\implies$ all diagonal elements $\neq 0$. We can rescale such that the diagonal elements are ± 1, hence a non-unique orthonormal basis. $\qquad\square$

Definition 3.22 (Timelike, Spacelike, Nulllike Vector)**:** Let $(\mathcal{M}, \mathbf{g})$ be a Lorentzian manifold, $V \in \mathcal{T}_p(\mathcal{M})$, $V \neq 0$. V is

- time-like iff $g(V,V) < 0$;

- null-like iff $g(V,V) = 0$;

- space-like iff $g(V,V) > 0$.

Definition 3.23 (Norm)**:** The norm of a space-like vector $V \in \mathcal{T}_p(\mathcal{M})$ is $|V| := \sqrt{g(V,V)}$.

Definition 3.24 (Angle)**:** The angle between space-like vectors $V,W \in \mathcal{T}_p(\mathcal{M})$ is $\theta = \cos^{-1} \frac{g(V,W)}{|V||W|}$.

3.3 Vector and Tensor Calculus on Manifolds

3.3.1 Covariant derivatives and connection

Now, the issue is, we cannot take the difference between vectors at diferent points. To be specific, $U \in T_p(M)$ and $V \in T_q(M)$ are two vectors that live in two different vector spaces. We thus need to define what we call - covariant derivative, on manifolds $\mathcal{M}$.

Definition 3.25 (Covariant Derivative for Functions)**:**

$$\nabla f : \ \mathcal{T}_p(\mathcal{M}) \to \mathbb{R}, \quad V \mapsto \nabla_V f := V(f) = V^\alpha \partial_\alpha f$$

∇f is a (0,1) tensor with $\nabla_\alpha f := \partial_\alpha f$

Remarks 3.12 (Gradient of scalar field)**.** *$\frac{\partial f}{\partial x^a}$ form components of a dual vector, which we call the gradient of a scalar field. We can always associate a vector to the gradient by $g^{ab}\partial_b f$. The covariant derivative is simply the gradient of the scalar field.*

Definition 3.26 (Covariant Derivative for Vector fields)**:**

$$\nabla V : \ T_p(\mathcal{M}) \to \mathcal{T}_p(\mathcal{M}), \quad X \mapsto \nabla_X V$$

which satisfies (f, g are functions and X, Y, V, W are vector fields):

- $\nabla_{fX+gY} V = f\nabla_X V + g\nabla_Y V$

- $\nabla_X (V + W) = \nabla_X V + \nabla_X W$

- $\nabla_X (fV) = f\nabla_X V + (\nabla_X f)V$ (Leibnitz Rule)

Equivalently, $\nabla V : \ \mathcal{T}_p^*(\mathcal{M}) \times \mathcal{T}_p(\mathcal{M}) \to \mathbb{R}, \quad (\eta, X) \mapsto \eta(\nabla_X V)$. ∇V is a (1,1) tensor with

$$V^\alpha_{\ ;\beta} = \nabla_\beta V^\alpha := \nabla V^\alpha_{\ \beta}$$

Definition 3.27 (Connection Coefficients)**:** Let $\{e^\mu\}$ be a basis of $\mathcal{T}_p(\mathcal{M})$, then we define the connection coefficients $\Gamma^\rho_{\mu\nu}$ to be

$$\nabla_\nu e_\mu := \nabla_{e_\nu} e_\mu = \Gamma^\rho_{\mu\nu} e_\rho$$

Remarks 3.13.

1. *In these notes, we only consider coordinate basis vectors $e_\mu = \partial_\mu$. This definition of the connection coefficients, however, is general and also holds for non-coordinate bases.*

2. *$\Gamma^\rho_{\mu\nu}$ are the expansion coefficients of $\nabla_\nu e_\mu$ in the basis $\{e_\rho\}$.*

3. *The word connection arises from the fact that we 'connect' the tangent spaces at different points $p, q \in \mathcal{M}$. Specifically, the connection coefficients give us the rate of change of the basis vector e_μ in the direction of the basis vector e_ν.*

4. *Here we use the convention that the second downstairs index of the connection, i.e. ν in denotes the direction in which the derivative is taken. The first index denotes the basis vector we are considering. In general relativity, the connection turns out to be symmetric in its downstairs indices, so that this convention does not really matter.*

Proposition 3.11 (Covariant derivative of a vector)**:**

$$\nabla_\nu W^\rho = \partial_\nu W^\rho + \Gamma^\rho_{\mu\nu} W^\mu$$

We get for $V = V^\mu e_\mu$ and $W = W^\mu e_\mu$. We thus have

$$
\begin{aligned}
\nabla_V W &= \nabla_V (W^\mu e_\mu) \\
&= V(W^\mu)e_\mu + W^\mu \nabla_V e_\mu \\
&= V^\nu e_\nu (W^\mu)e_\mu + W^\mu \nabla_{V^\mu e_\nu} e_\mu \\
&= V^\nu e_\nu (W^\mu)e_\mu + W^\mu V^\nu \nabla_\nu e_\mu \\
&= V^\mu (\partial_\nu W^\rho + W^\mu \Gamma^\rho_{\mu\nu})e_\rho \\
\implies (\nabla_V W)^\rho &= V^\mu \partial_\nu W^\rho + V^\nu \Gamma^\rho_{\mu\nu} W^\mu
\end{aligned}
$$

but since V is arbitrary, we take $\nabla_\nu W^\rho$, which is the desired result $\partial_\nu W^\rho + \Gamma^\rho_{\mu\nu} W^\mu$. We thus denote this as $W^\rho_{\ ;\nu}$.

Proposition 3.12 (Transform connection)**:** Under coordinate transformation $x^\mu \to \tilde{x}^\alpha$, the Connection coefficient transform as

$$\tilde{\Gamma}^\sigma_{\mu\nu} = \frac{\partial \tilde{x}^\sigma}{\partial x^\rho} \frac{\partial x^\alpha}{\partial \tilde{x}^\nu} \frac{\partial x^\beta}{\partial \tilde{x}^\mu} \Gamma^\rho_{\beta\alpha} + \frac{\partial \tilde{x}^\sigma}{\partial x^\rho} \frac{\partial^2 x^\rho}{\partial \tilde{x}^\nu \partial \tilde{x}^\mu}$$

Proof. With $\tilde{\Gamma}^\sigma_{\mu\nu} \tilde{\partial}_\sigma = \nabla_{\tilde{\partial}_\nu} \tilde{\partial}_\mu$ in mind, we compute

$$\begin{aligned}
\tilde{\Gamma}^\sigma_{\mu\nu} \tilde{\partial}_\sigma &= \frac{\partial x^\alpha}{\partial \tilde{x}^\nu} \nabla_{\partial_\alpha} \left(\frac{\partial x^\beta}{\partial \tilde{x}^\mu} \partial_\beta \right) \\
&= \frac{\partial x^\alpha}{\partial \tilde{x}^\nu} \frac{\partial^2 x^\beta}{\partial x^\alpha \partial \tilde{x}^\mu} \partial_\beta + \frac{\partial x^\alpha}{\partial \tilde{x}^\nu} \frac{\partial x^\beta}{\partial \tilde{x}^\nu} \nabla_{\partial_\alpha} (\partial_\beta) \\
&= \frac{\partial^2 x^\beta}{\partial \tilde{x}^\nu \partial \tilde{x}^\mu} \partial_\beta + \frac{\partial x^\alpha}{\partial \tilde{x}^\nu} \frac{\partial x^\beta}{\partial \tilde{x}^\mu} \Gamma^\rho_{\beta\alpha} \partial_\rho \\
&= \left[\frac{\partial^2 x^\beta}{\partial \tilde{x}^\nu \partial \tilde{x}^\mu} + \frac{\partial x^\alpha}{\partial \tilde{x}^\nu} \frac{\partial x^\beta}{\partial \tilde{x}^\mu} \Gamma^\rho_{\beta\alpha} \right] \frac{\partial \tilde{x}^\sigma}{\partial x^\rho} \tilde{\partial}_\sigma
\end{aligned}$$

where $\partial_\rho = \frac{\partial \tilde{x}^\sigma}{\partial x^\rho} \tilde{\partial}_\sigma$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Remarks 3.14. *So, $\Gamma^\sigma_{\mu\nu}$ is not a tensor, but the difference of two connections is, i.e. Torsion tensor. Another example of what looks like a tensor, but actually isn't: $\partial_\nu W^\mu$.*

Definition 3.28 (Torsion Tensor)**:**

$$T_{\mu\nu}{}^\lambda :=:= \Gamma^\lambda_{\mu\nu} - \Gamma^\lambda_{\nu\mu}$$

Definition 3.29 (Torsion Free)**:** The connection Γ is called torsion free iff $T_{\mu\nu}{}^\lambda = 0$.

Proposition 3.13 (Covariant Derivative of Tensor)**:** For a (r, s) tensor, its covariant derivative is

$$\begin{aligned}
\nabla_\rho T^{\mu_1 \cdots \mu_r}{}_{\nu_1 \ldots \nu_s} &= \partial_\rho T^{\mu_1 \cdots \mu_r}{}_{\nu_1 \ldots \nu_s} \\
&+ \Gamma^{\mu_1}_{\sigma\rho} T^{\sigma \mu_2 \cdots \mu_r}{}_{\nu_1 \ldots \nu_s} + \ldots + \Gamma^{\mu_r}_{\sigma\rho} T^{\mu_1 \cdots \mu_{r-1} \sigma}{}_{\nu_1 \ldots \nu_s} \\
&- \Gamma^\sigma_{\nu_1 \rho} T^{\mu_1 \cdots \mu_r}{}_{\sigma \nu_2 \ldots \nu_s} - \ldots - \Gamma^\sigma_{\nu_s \rho} T^{\mu_1 \cdots \mu_r}{}_{\nu_1 \ldots \nu_{s-1} \sigma}
\end{aligned}$$

This can be obtained from Leibniz rule, where for (r, s) tensor T, ∇T is of rank $(r, s + 1)$.

Example 3.8 (Covariant derivative of a one-form)**.**

$$\nabla_V (\eta(W)) := (\nabla_V \eta)(W) + \eta(\nabla_V W) \implies (\nabla_V \eta)(W) = \nabla_V (\eta(W)) - \eta(\nabla_V W)$$

So, $\nabla\eta$ is a (0,2) tensor since

$$\begin{aligned}
(\nabla \eta)(V, W) = (\nabla_V \eta)(W) &= \nabla_V (\eta_\mu W^\mu) - \eta_\mu (\nabla_V W)^\mu \\
&= V^\rho \partial_\rho (\eta_\mu W^\mu) - \eta_\mu (V^\rho \partial_\rho W^\mu + V^\rho \Gamma^\mu_{\nu\rho} W^\nu) \\
&= V^\rho W^\mu \partial_\rho \eta_\mu - \Gamma^\mu_{\nu\rho} \eta_\mu V^\rho W^\nu \\
&= (\partial_\rho \eta_\mu - \Gamma^\nu_{\mu\rho} \eta_\nu) V^\rho W^\mu
\end{aligned}$$

So the components of $\nabla_\rho \eta_\mu := (\nabla\eta)_{\rho\mu}$ (also written as $\eta_{\mu;\rho}$ is $\partial_\rho \eta_\mu - \Gamma^\nu_{\mu\rho} \eta_\nu$).

Example 3.9 (Covariant derivative of rank-two tensor)**.**

$$\nabla_c T^{ab} = \partial_c T^{ab} + \Gamma^a_{cd} T^{db} + \Gamma^b_{cd} T^{ad}$$

Corollary 3.1: The covariant derivative of a vector field should be a tensor.

By construction of covariant derivative, this must be true. Let's verify.

Proof. Starting from the definition of a covariant derivative, and taking the result of Proposition 3.12 and the tensor transformation laws:

$$\tilde{\nabla}_\nu \tilde{W}^\rho = \tilde{\partial}_\nu \tilde{W}^\rho + \tilde{\Gamma}^\rho_{\mu\nu} \tilde{W}^\mu$$

$$= \frac{\partial x^\alpha}{\partial \tilde{x}^\nu} \frac{\partial}{\partial x^\alpha} \left(\frac{\partial \tilde{x}^\rho}{\partial x^\beta} W^\beta \right) + \frac{\partial x^\alpha}{\partial \tilde{x}^\mu} \frac{\partial x^\beta}{\partial \tilde{x}^\nu} \frac{\partial \tilde{x}^\rho}{\partial x^\gamma} \Gamma^\gamma_{\alpha\beta} \frac{\partial \tilde{x}^\mu}{\partial x^\delta} W^\delta + \frac{\partial \tilde{x}^\rho}{\partial x^\lambda} \tilde{\partial}_\nu \left(\frac{\partial x^\lambda}{\partial \tilde{x}^\mu} \right) \frac{\partial \tilde{x}^\mu}{\partial x^\beta} W^\beta$$

$$= \frac{\partial x^\alpha}{\partial \tilde{x}^\nu} \frac{\partial \tilde{x}^\rho}{\partial x^\beta} \partial_\alpha W^\beta + W^\beta \tilde{\partial}_\nu \left(\frac{\partial \tilde{x}^\rho}{\partial x^\beta} \right) + \frac{\partial x^\beta}{\partial \tilde{x}^\nu} \frac{\partial \tilde{x}^\rho}{\partial x^\gamma} \Gamma^\gamma_{\alpha\beta} W^\alpha + W^\beta \frac{\partial \tilde{x}^\rho}{\partial x^\lambda} \frac{\partial \tilde{x}^\mu}{\partial x^\beta} \tilde{\partial}_\nu \left(\frac{\partial x^\lambda}{\partial \tilde{x}^\mu} \right)$$

but $\frac{\partial \tilde{x}^\mu}{\partial x^\beta} \frac{\partial x^\lambda}{\partial \tilde{x}^\mu} = \delta^\lambda{}_\beta$, which means

$$\frac{\partial \tilde{x}^\mu}{\partial x^\beta} \tilde{\partial}_\nu \left(\frac{\partial x^\lambda}{\partial \tilde{x}^\mu} \right) = -\frac{\partial x^\lambda}{\partial \tilde{x}^\mu} \tilde{\partial}_\nu \left(\frac{\partial \tilde{x}^\mu}{\partial x^\beta} \right) \implies W^\beta \frac{\partial \tilde{x}^\rho}{\partial x^\lambda} \frac{\partial \tilde{x}^\mu}{\partial x^\beta} \tilde{\partial}_\nu \left(\frac{\partial x^\lambda}{\partial \tilde{x}^\mu} \right) = -W^\beta \delta^\rho{}_\mu \tilde{\partial}_\nu \left(\frac{\partial \tilde{x}^\mu}{\partial x^\beta} \right) = -W^\beta \tilde{\partial}_\nu \left(\frac{\partial \tilde{x}^\rho}{\partial x^\beta} \right)$$

Hence, by definition of covariant derivative again

$$\tilde{\nabla}_\nu \tilde{W}^\rho = \frac{\partial x^\alpha}{\partial \tilde{x}^\nu} \frac{\partial \tilde{x}^\rho}{\partial x^\beta} \partial_\alpha W^\beta + \frac{\partial x^\beta}{\partial \tilde{x}^\nu} \frac{\partial \tilde{x}^\rho}{\partial x^\gamma} \Gamma^\gamma_{\alpha\beta} W^\alpha = \frac{\partial x^\alpha}{\partial \tilde{x}^\nu} \frac{\partial \tilde{x}^\rho}{\partial x^\beta} \nabla_\alpha W^\beta$$

$\square$

3.3.2 The Levi-Civita Connection

In general manifolds, there is no more fundamental structure on the manifold that determines the connection, so it is part of defining the geometry to equip it with a connection of your choice. Note that it is not even necessary to have a metric on the manifold. Although we need no metric for the connection, a metric singles out a special connection (due to the fundamental theorem of Riemannian geometry). We will see later that this metric is the Riemann metric. The Levi-Civita Connection is used to find the connection coefficients.

Definition 3.30 (Christoffel symbols)**:** We define the Christoffel symbols to be short for

$$\begin{Bmatrix} \beta \\ \mu \ \nu \end{Bmatrix} := \frac{1}{2} g^{\beta\rho} (\partial_\mu g_{\nu\rho} + \partial_\nu g_{\rho\mu} - \partial_\rho g_{\mu\nu})$$

As we will see, the Christoffel symbols are actually a connection. The second downstairs index will be conventionally our derivative index.

Theorem 3.4: On a manifold $\mathcal{M}$ with metric g, $\exists$ a unique connection with

- $\Gamma^\alpha_{\mu\nu} = \Gamma^\alpha_{\nu\mu} = \begin{Bmatrix} \alpha \\ \mu \ \nu \end{Bmatrix}$, i.e. Christoffel symbols are torsion free;

- $\nabla g = 0$, i.e. 'metric compatible'.

This connection is the Levi-Civita connection.

Proof. ($\rightarrow$) Let $\Gamma^\alpha_{\beta\gamma}$ be 'metric compatible', and symmetric. then

$$\nabla_\alpha g_{\beta\gamma} = 0 \implies \partial_\alpha g_{\beta\gamma} = \Gamma^\rho_{\beta\alpha} g_{\rho\gamma} + \Gamma^\rho_{\gamma\alpha} g_{\beta\rho}$$

From the definition of the Christoffel symbols,

$$\begin{Bmatrix} \mu \\ \beta \ \gamma \end{Bmatrix} = \frac{1}{2} g^{\mu\nu} (\partial_\beta g_{\gamma\nu} + \partial_\gamma g_{\nu\beta} - \partial_\nu g_{\beta\gamma})$$

$$= \frac{1}{2} g^{\mu\nu} (\Gamma^\rho_{\gamma\beta} g_{\rho\nu} + \Gamma^\rho_{\nu\beta} g_{\gamma\rho} + \Gamma^\rho_{\nu\gamma} g_{\rho\beta} + \Gamma^\rho_{\beta\gamma} g_{\nu\rho} - \Gamma^\rho_{\beta\nu} g_{\rho\gamma} - \Gamma^\rho_{\gamma\nu} g_{\beta\rho})$$

$$= \Gamma^\mu_{\beta\gamma} \tag{1}$$

($\leftarrow$) Likewise, $\Gamma^\alpha_{\beta\gamma} = \begin{Bmatrix} \alpha \\ \mu \ \nu \end{Bmatrix}$, then we can check that is symmetric in β and γ and so

$$\nabla_\alpha g_{\beta\gamma} = \partial_\alpha g_{\beta\gamma} - \Gamma^\rho_{\beta\alpha} g_{\rho\gamma} - \Gamma^\rho_{\gamma\alpha} g_{\beta\rho}$$

$$= \partial_\alpha g_{\beta\gamma} - \frac{1}{2} g^{\rho\sigma} [(\partial_\beta g_{\alpha\sigma} + \partial_\alpha g_{\sigma\beta} - \partial_\sigma g_{\beta\alpha}) g_{\rho\gamma} + (\partial_\gamma g_{\alpha\sigma} + \partial_\alpha g_{\sigma\gamma} - \partial_\sigma g_{\gamma\alpha}) g_{\rho\beta}]$$

$$= \partial_\alpha g_{\beta\gamma} - \frac{1}{2} [(\partial_\beta g_{\alpha\sigma} + \partial_\alpha g_{\sigma\beta} - \partial_\sigma g_{\beta\alpha}) + (\partial_\gamma g_{\alpha\sigma} + \partial_\alpha g_{\sigma\gamma} - \partial_\sigma g_{\gamma\alpha})]$$

$$= \partial_\alpha g_{\beta\gamma} - \frac{1}{2} (\partial_\alpha g_{\gamma\beta} + \partial_\alpha g_{\beta\gamma}) = 0$$

□

Remarks 3.15.

1. *In GR, we generally assume the connection $\Gamma^\alpha_{\beta\gamma}$ to be the Levi-Civita one.*

2. *Torsion free connection would also mean commutative action on scalar fields.*

3. *When the metric compatibility is satisfied, we can interchange the order of raising/lowering indices and covariant differentiation. For instance,*

$$\nabla_c T^{ab} = \nabla_c (g^{bd} T^a{}_d = (\nabla_c g^{bd}) T^a{}_d + g^{bd}(\nabla_c T^a{}_d) = g^{bd}(\nabla_c T^a{}_d)$$

   Hence, the index associated with a covariant derivative is a genuine tensor index.

4. *We sometimes require the connection coefficients summed over the upper and a lower index. Since $\nabla_c g_{ab} = 0$, we have*

$$\partial_c g_{ab} = \Gamma^d_{ca} g_{db} + \Gamma^d_{cb} g_{ad} \implies g^{ab}\partial_c g_{ab} = g^{ab}(\Gamma^d_{ca} g_{db} + \Gamma^d_{cb} g_{ad}) = 2g^{ab}g_{db}\Gamma^d_{ca}) = 2\Gamma^a_{ac}$$

   This follows from $(\det M)^{-1}\partial_c \det M = \mathrm{Tr}(M^{-1}\partial_c M)$ for some invertible matrix M. We then have $\Gamma^a_{ac} = \frac{1}{\sqrt{|g|}}\partial_c\sqrt{|g|}$.

5. *The covariant derivative constructed with the metric connection has the nice property that it reduces to partial differentiation in local Cartesian coordinate. This follows from the vanishing derivative of the metric at $p \in \mathcal{M}$, $\frac{\partial g_{ab}}{\partial x^c}|_p = 0$. This is very important for enforcing the equivalence principle as it is straightforward to check that some law of physics, written as a tensor equation, reduces to its usual special-relativistic form in local Cartesian coordinates. Conversely, we can easy formulate physical theories in general spacetime by promoting derivatives in Euclidean space to covariant derivatives.*

6. *Divergence, curl and Laplacian with Levi-Civita connection:*

3.3.3 Parallel transport

A connection gives us the notion of a tensor that does not change along a curve.

Definition 3.31 (Integral Curve)**:** An integral curve λ of a vector field V through $p \in \mathcal{M}$ is a curve through p and whose tangent at every point q along the curve is V_q.

In coordinates,

$$\frac{d}{dt}\Big|_\lambda = V \implies \frac{d}{dt}x^\mu(\lambda(t)) = V^\mu(x^\alpha)$$

with $x^\mu(\lambda(t_0)) = x^\mu(p)$ has a unique solution by ODE theory.

Definition 3.32 (Parallel Transport)**:** Let V be a vector field and $\mathcal{C}$ be an integral curve. A tensor T is parallel transported along $\mathcal{C}$ if $\nabla_V T = 0$ along $\mathcal{C}$.

Proposition 3.14: The tangent of an affinely parametrized geodesic is parallel transported along itself.

Proof.

$$U^\mu \nabla_\mu U^\alpha = U^\mu \partial_\mu U^\alpha + U^\mu \Gamma^\alpha_{\nu\mu} U^\nu = \frac{dx^\mu}{d\lambda}\frac{\partial}{\partial x^\mu}\frac{dx^\alpha}{d\lambda} + \frac{dx^\mu}{d\lambda}\Gamma^\alpha_{\nu\mu}\frac{dx^\nu}{d\lambda} = 0$$

This leads to $\frac{d^2 x^\alpha}{d\lambda^2} + \Gamma^\alpha_{\nu\mu}\frac{dx^\nu}{d\lambda}\frac{dx^\mu}{d\lambda} = 0$ with non-affine parameter $\nabla_U U = fU$. $\qquad\square$

Proposition 3.15: $\nabla_V T = 0$ determines T uniquely along the curve of transport.

Proof. In coordinates x^μ, the curve is $x^\mu(\lambda)$, then

$$V^\sigma \nabla_\sigma \delta^\mu{}_\nu = V^\sigma \partial_\sigma \delta^\mu{}_\nu + \Gamma^\mu_{\rho\sigma} T^\rho{}_\nu V^\sigma - \Gamma^\rho_{\nu\sigma} T^\mu{}_\rho V^\sigma = \frac{d}{d\lambda}T^\mu{}_\nu + \Gamma^\mu_{\rho\sigma} T^\rho{}_\mu V^\sigma - \Gamma^\rho_{\nu\sigma} \delta^\mu{}_\rho V^\sigma = 0$$

ODE theory thus guarantees a solution $\forall$ components of T. $\qquad\square$

Proposition 3.16: Parallel transport T along a curve from p to q provides an isomorphism between tensors at p and q.

Unlike Special Relativity, this is path dependent in General Relativity.

Proposition 3.17: Parallel transport preserves the length of vectors.

Proof.

$$\frac{d}{d\lambda}(W_\alpha W^\alpha) = V^\mu \nabla_\mu (W_\alpha W^\alpha) = 2W_\alpha V^\mu \nabla_\mu W^\alpha$$

but $V^\mu \nabla W^\alpha = 0$. $\qquad\square$

As such, geodesics do not change their timelike, spacelike or null character.

Definition 3.33 (Acceleration Along Timelike Curve)**:** We define the acceleration a^μ along a time-like curve by $a^\mu := U^\rho \nabla_\rho U^\alpha$.

Proposition 3.18: This curve is a geodesic if $a^\mu = 0$ or $a^\mu = fU^\mu$ such that f is a function.

We can also define $\frac{D}{D\lambda} := \frac{dx^\rho}{d\lambda}\nabla_\rho$ such that the parallel transport of a tensor is $\frac{DT}{D\lambda} = 0$.

3.3.4 Geodesics

Definition 3.34 (Timelike, Spacelike, Nulllike Curve)**:** A curve $\mathcal{C}$ is time-like (null,spacelike) at a point $p \in \mathcal{M}$ iff its tangent vector at that point is time-like (null, spacelike).

This property can change along the curve parametrized by $\lambda(t)$.

Definition 3.35 (Proper Length, Time)**:** The proper length along a space-like curve is $\lambda_{space}(t)$

$$s := \int_{t_0}^{t_1} \sqrt{g(V,V)|_{\lambda_{space}(t)}} dt = \int_{t_0}^{t_1} \sqrt{g_{\alpha\beta} \frac{dx^\alpha}{dt} \frac{dx^\beta}{dt}} dt$$

The proper time along a time-like curve is $\lambda_{time}(t)$

$$\tau(t_1) := \int_{t_0}^{t_1} \sqrt{-g(V,V)|_{\lambda_{space}(t)}} dt = \int_{t_0}^{t_1} \sqrt{-g_{\alpha\beta} \frac{dx^\alpha}{dt} \frac{dx^\beta}{dt}} dt$$

Definition 3.36 (Four-Velocity)**:** The four-velocity along time-like curves $u^\mu := \frac{dx^\mu}{dt}|_{\lambda(\tau)}$.

Proposition 3.19: $g_{\mu\nu} u^\mu u^\nu = -1$

Theorem 3.5 (Noether's Theorem)**:** The action $\mathcal{S} = \int \mathcal{L}(q_k, \dot{q}_k, \lambda) d\lambda$ is extremized by the curve satisfying the Euler-Lagrange equations

$$\frac{d}{d\lambda} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_k} \right) = \frac{\partial \mathcal{L}}{\partial q_k}$$

Corollary 3.2: Noether's theorem have a few simplifications:

- If $\mathcal{L}$ not explicitly dependent on $q_k \implies p_k := \frac{\partial \mathcal{L}}{\partial \dot{q}_k}$ is conserved along the curve that extremizes $\mathcal{S}$;

- If $\mathcal{L}$ is not explicitly dependent on the parameter $\lambda \implies I := \dot{q}_k \frac{\partial \mathcal{L}}{\partial \dot{q}_k} - \mathcal{L}$ is conserved along the curve.

Proof. Review IB Variational Principles for the above Theorem and Corollary. $\qquad \square$

Definition 3.37 (Christoffel symbols)**:** We define the Christoffel symbols to be short for

$$\left\{ \begin{matrix} \beta \\ \mu \ \nu \end{matrix} \right\} := \frac{1}{2} g^{\beta\rho}(\partial_\mu g_{\nu\rho} + \partial_\nu g_{\rho\mu} - \partial_\rho g_{\mu\nu})$$

Proposition 3.20 (Geodesic Equation - Version 1)**:** Consider time-like curves from A ($\lambda = 0$) and B ($\lambda = 1$). The action is $\mathcal{S} = \int_0^1 \mathcal{L} d\lambda$ with $\mathcal{L} = \sqrt{-g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu}$, then after extremizing the action, we will obtain one possible version of the Geodesic Equation:

$$\frac{d^2 x^\beta}{d\tau^2} + \left\{ \begin{matrix} \beta \\ \mu \ \nu \end{matrix} \right\} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = 0 \qquad\qquad (\dagger)$$

Here, $\mathcal{S}$ is invariant under the reparametrization $\kappa(\lambda)$, where $\frac{d\kappa}{d\lambda} > 0$, i.e.

$$\mathcal{S} = \int_0^1 \sqrt{-g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} d\lambda = \int_{\kappa(0)}^{\kappa(1)} \sqrt{-g_{\mu\nu} \frac{dx^\mu}{d\kappa} \frac{dx^\nu}{d\kappa}} d\kappa$$

Proof. The corresponding Euler-Lagrange equations for $\mathcal{S}$ will be

$$\frac{d}{d\lambda} \left(\frac{1}{2\mathcal{L}} (-g_{\mu\nu} \delta^\mu_\alpha \dot{x}^\nu - g_{\mu\nu} \dot{x}^\mu \delta^\nu_\alpha) \right) - \frac{1}{2\mathcal{L}} (-\dot{x}^\mu \dot{x}^\nu \partial_\alpha g_{\mu\nu}) = \frac{d}{d\lambda} \left(-\frac{g_{\mu\alpha} \dot{x}^\mu}{\mathcal{L}} \right) + \frac{\dot{x}^\mu \dot{x}^\nu \partial_\alpha g_{\mu\nu}}{2\mathcal{L}} = 0$$

We change the parameter from λ to $\tau(\lambda) = \int_0^\lambda \sqrt{-g_{\mu\nu} \frac{dx^\mu}{d\tilde{\lambda}} \frac{dx^\nu}{d\tilde{\lambda}}} d\tilde{\lambda}$ (proper-time) where $\frac{d\tau}{d\lambda} = \mathcal{L}$, then

$$-\mathcal{L} \frac{d}{d\tau} \left(g_{\mu\alpha} \frac{dx^\mu}{d\tau} \right) + \frac{\mathcal{L}}{2} \frac{dx^\mu}{d\tau} \frac{dx^\mu}{d\tau} \partial_\alpha g_{\mu\nu} = 0$$

Take dot product with $g^{\beta\alpha}$ gives

$$\frac{d^2 x^\mu}{d\tau^2} g_{\mu\alpha} + \partial_\nu g_{\mu\alpha} \frac{dx^\nu}{d\tau} \frac{dx^\mu}{d\tau} - \frac{1}{2} \partial_\alpha g_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = 0$$

which gives us the desired Geodesic Equation. $\qquad \square$

For space-like geodesics, we take $\tilde{\mathcal{L}} = \sqrt{g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu}$ such that $\frac{ds}{d\lambda} = \tilde{\mathcal{L}}$. In that case, the Geodesic Equation will have s as the parameter instead.

Proposition 3.21 (Geodesic Equation - Version 2): Consider a different variation, where we do not restrict ourselves to time-like geodesics. The action is now $\hat{\mathcal{S}} = \int_A^B \hat{\mathcal{L}} d\lambda$ with the new Lagrangian $\hat{\mathcal{L}} = g_{\alpha\beta} \frac{dx^\alpha}{d\lambda} \frac{dx^\beta}{d\lambda}$. Then after extremizing the action, we will obtain another version of the Geodesic Equation:

$$\ddot{x}^\alpha + \begin{Bmatrix} \alpha \\ \nu\ \beta \end{Bmatrix} \dot{x}^\nu \dot{x}^\beta = 0 \tag{*}$$

where $\dot{\ }$ represents $\frac{d}{d\lambda}$ such that λ parametrizes the curve.

While it is great that we do not need to restrict ourselves to time-like curves, it is however not invariant under reparametrization.

Proof. Working with the Euler-Lagrange equations gives us our desired Geodesic Equation. $\qquad\square$

So far so good. But now take (†) and let $\tau = \tau(\lambda)$ such that $\frac{d\tau}{d\lambda} > 0$. We have $\frac{d}{d\tau} = \frac{d\lambda}{d\tau}\frac{d}{d\lambda}$ and so $\frac{d^2}{d\tau^2} = \frac{d^2\lambda}{d\tau^2}\frac{d}{d\lambda} + (\frac{d\lambda}{d\tau})^2 \frac{d^2}{d\lambda^2}$. We thus obtained

$$\frac{d^2 x^\alpha}{d\lambda^2} + \begin{Bmatrix} \alpha \\ \nu\ \beta \end{Bmatrix} \frac{dx^\nu}{d\lambda}\frac{dx^\beta}{d\lambda} = -\left(\frac{d\lambda}{d\tau}\right)^{-2} \frac{d^2\lambda}{d\tau^2}\frac{dx^\alpha}{d\lambda} \tag{**}$$

The form of the geodesic equation is no longer the same. We thus demonstrate that $\hat{\mathcal{S}}$ is not invariant under a change of parametrization. Different variation gives different curves! Both equations (*) and (†) only agree if $\frac{d^2\lambda}{d\tau^2} = 0$. Equivalently, λ is linear in τ.

Definition 3.38 (Affine Parameter): The parameter λ along a timelike (spacelike) curve is affine iff it is linearly related to the proper time τ (proper distance s). For a non-affine parameter, the corresponding Geodesic Equation is (**)

In summary,

Proposition 3.22: If a curve $\mathcal{C} : I \subset \mathbb{R} \to \mathcal{M}, \ \lambda \mapsto x^\alpha(\lambda)$, then if it

- satisfies (*) $\implies \mathcal{C}$ is a geodesic and λ is affine;

- satisfies (**) with non-zero right hand side $\implies \mathcal{C}$ is a geodesic but λ is not affine;

- satisfies neither (*) nor (**) $\implies \mathcal{C}$ is not a geodesic.

From ODE theory, the solutions to (*) and (**) are unique if x^α and $\dot{x}^\alpha$ fixed at $\lambda = \lambda_0$.

Postulate 3.1 (Geodesic Postulate). *Test particles with positive (zero) rest mass move on time-like (null) geodesics.*

$\hat{\mathcal{L}}$ gives an easy way to calculate the Christoffel symbols.

Example 3.10 (Christoffel Symbols for Schwarzschild Metric). *The Schwarzschild metric (as we will see later) in Schwarzschild coordinates is*

$$ds^2 = -f dt^2 + f^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2$$

with $f = 1 - \frac{2M}{r}$ where M is a constant. We have

$$\hat{\mathcal{L}} = f\dot{t}^2 - f^{-1}\dot{r}^2 - r^2\dot{\theta}^2 - r^2\sin^2\theta\dot{\phi}^2$$

The Euler-Lagrange for $t(\tau)$ is $\frac{d}{d\tau}(2f\dot{t}) = 0 \implies \frac{d^2\tau}{d\tau^2} + f^{-1}\frac{df}{dr}\dot{t}\dot{r} = 0$ and so

$$\begin{Bmatrix} t \\ t\ r \end{Bmatrix} = \begin{Bmatrix} t \\ r\ t \end{Bmatrix} = \frac{df/dr}{2f} = \frac{1}{1 - (2M/r)}\frac{M}{r^2}$$

and $\begin{Bmatrix} t \\ \nu\ \mu \end{Bmatrix} = 0$ otherwise for $\mu, \nu = \theta, \phi$.

4 More on Special Relativity

4.1 Minkowski Spacetime and Particle Dynamics

4.2   Electromagnetism

27

5   Spacetime Curvature

6   Gravitational Field Equations

7   Schwarzschild Solution

8   Classic Tests of General Relativity

31

9   Cosmology

32